ISSN: 2663-5798 || Arab Journal for Scientific Publishing

الإصدار الثامن – العدد الثالث والثمانون
تاريخ الإصدار: 2 – ايلول – 2025م
www.ajsp.net

"Prehospital Triage Tools in Disaster Situations: A Systematic Review of Accuracy and Effectiveness"

Researchers:

Abdulaziz Nafea T Alsulami[1], Fawaz Othman A Almasri[1], Saleh Mohammed Bajhzer[2], Abdullah Ismail M Albalosh[2], Alharbi Mamdouh Saleh[1], Alzahrani, Saeed Atiah A[1], Althubyani, Hassan Mohsen S[1], Wajdy Hassan Alharbi[1]

1. Emergency Medical Technician, Saudi Red Crescent Authority
2. Emergency Medical Specialist, Saudi Red Crescent Authority

ISSN: 2663-5798 || Arab Journal for Scientific Publishing

الإصدار الثامن – العدد الثالث والثمانون
تاريخ الإصدار: 2 – ايلول – 2025م
www.ajsp.net

Abstract:

Background: Disasters and mass-casualty incidents (MCIs) require rapid, defensible prehospital triage to prioritize scarce resources, yet the accuracy and effectiveness of commonly used tools remain debated. **Objective:** To synthesize contemporary evidence on the diagnostic accuracy and operational performance of prehospital triage tools used in disaster/MCI settings, with attention to adult and pediatric applications. **Methods:** We conducted a systematic review following PRISMA-DTA principles, targeting studies from 2000 onward that evaluated START, SALT, CareFlight, Triage Sieve/Sort, Sacco, JumpSTART, PTT/SPTT, MPTT/MPTT-24, MITT, and Ten-Second Triage. Eligible designs included prospective/retrospective cohorts, registry re-triage, simulation/live exercises, and guideline-linked implementation studies. Primary outcomes were sensitivity/specificity for identifying need for life-saving interventions (LSIs) and under-/over-triage; secondary outcomes included mortality, major injury, time-to-triage, inter-rater reliability, and usability. **Results:** Evidence indicates that legacy algorithms (e.g., START, SALT) show wide and often insufficient diagnostic accuracy, with under- and over-triage rates that frequently exceed accepted trauma-system benchmarks. In contrast, physiology-based tools (MPTT/MPTT-24) demonstrate higher sensitivity for LSI and have informed adoption of the Major Incident Triage Tool (MITT) in national guidance. Pediatric performance remains the greatest gap: data for JumpSTART are limited and mixed, whereas SPTT and age-inclusive modern tools (e.g., MPTT-24/MITT) show improved discrimination but require prospective validation in real incidents. Training innovations (e.g., XR/VR) and MUCC-aligned frameworks may enhance consistency and interoperability but need stronger effectiveness evidence. **Conclusions:** Contemporary data favor LSI-anchored, physiology-based triage for primary disaster use, with careful system-level integration, routine training, and ongoing audit. Pediatric triage warrants prioritized research and cautious adoption of higher-sensitivity tools pending real-world validation.

Keywords: **Prehospital, Triage Tools, Disaster Situations, Accuracy, Effectiveness**

---

3. Emergency Medical Technician, Saudi Red Crescent Authority
4. Emergency Medical Specialist, Saudi Red Crescent Authority

Introduction

Disasters and mass-casualty incidents (MCIs) abruptly overwhelm local resources and demand rapid, defensible methods to prioritize scarce prehospital care. Triage—the structured sorting of many casualties to maximize population benefit—remains the cornerstone of early MCI management, yet persistent variation in tools and performance has fueled debate about which algorithms are most accurate and effective in the field (Zhang, et al, 2023). In practice, prehospital triage seeks a pragmatic balance: it must be fast and simple enough for cross-disciplinary responders to apply under duress, while still reliably identifying patients who need immediate, life-saving interventions (LSIs) and safe deferral of lower-acuity care.

Over the last two decades, multiple adult and pediatric MCI triage systems have been proposed or widely adopted—including START (Simple Triage and Rapid Treatment), SALT (Sort, Assess, Lifesaving Interventions, Treatment/Transport), CareFlight, Triage Sieve/Sort, the Sacco Triage Method, JumpSTART (pediatric), and more recent physiology-based tools such as the Modified Physiological Triage Tool (MPTT/MPTT-24) and the UK's Major Incident Triage Tool (MITT). To harmonize this heterogeneity, the United States introduced the Model Uniform Core Criteria (MUCC), a science- and consensus-based national guideline that specifies 24 essential elements for any mass-casualty triage system, with SALT developed as MUCC-compliant guidance for broad interoperability.

Despite widespread training and use, the empirical accuracy of several legacy algorithms is increasingly questioned. A 2022 meta-analysis concluded that START lacks sufficient diagnostic accuracy to serve as a reliable disaster triage tool, echoing earlier concerns about inconsistent performance across scenarios (Franc, et al. 2022). Comparative reviews similarly report wide accuracy ranges for START and mixed findings for SALT, with both systems exhibiting non-trivial under- and over-triage rates—an operationally critical issue because under-triage risks preventable mortality, whereas over-triage consumes scarce resources (Purwadi, et al., 2021).

Benchmarks used in trauma systems (e.g., American College of Surgeons Committee on Trauma targets of <5% under-triage and ~25–35% over-triage) are frequently exceeded in field studies, underscoring the need for more discriminating prehospital tools during MCIs (Biesboer, et al., 2023).

Contemporary physiology-based triage has evolved in response. The MPTT and its refinement (MPTT-24) demonstrated superior sensitivity for identifying patients requiring LSIs compared with legacy UK tools (Sieve/Sort), informing adoption of the unified MITT across NHS services beginning in the first of 21 centuries (Vassallo, et al, 2021). This move reflects a broader shift toward tools optimized for high sensitivity to time-critical injury, explicit links to an LSI "gold standard," and applicability across the age spectrum. Early evaluations of MITT emphasize speed, reproducibility, and cross-professional usability—attributes crucial to prehospital implementation under austere conditions (Vassallo, etal., 2022).

Pediatric disaster triage presents additional challenges. Evidence for JumpSTART is limited and mixed: one systematic review reported a pooled accuracy around 57% with relatively low over-triage, while other comparative analyses suggest pediatric-specific tools (e.g., SPTT) or age-inclusive modern tools (e.g., MITT) may improve identification of high-acuity children (Stéfani, et al., 2023). These uncertainties matter operationally, given the known vulnerability of special populations and the ethical imperative to minimize preventable harm during MCIs.

Methodological constraints complicate the evidence base. Much of the triage literature relies on simulations, registry re-triage, or tabletop exercises rather than prospective data from real MCIs; reference standards vary (e.g., mortality, ISS thresholds, need for LSIs), and outcome definitions are not uniform across studies. Systematic reviews of prehospital triage tools for major trauma highlight substantial variability in sensitivity/specificity and heterogeneous comparators limitations that likely generalize, and sometimes amplify, in disaster contexts (Gianola, et al., 2021). Nevertheless, converging data indicate that accuracy deficits remain common across systems and settings, with significant implications for survival and system efficiency.

At the same time, implementation science and technology are reshaping the landscape. National guidance (e.g., MUCC) promotes interoperability and standardization, while digital decision support and immersive VR/simulation platforms are being explored to enhance training, consistency, and real-time performance of triage at scale .EMS.govBioMed Centraljmir.org These developments reinforce the need to continuously appraise the accuracy and effectiveness of prehospital triage tools—not only in controlled environments but, as feasible, in real incidents and large multi-agency exercises.

Accordingly, this systematic review synthesizes contemporary evidence on the diagnostic accuracy and operational effectiveness of prehospital triage tools used in disaster and MCI settings. We focus on (a) discrimination for LSIs and other clinically meaningful outcomes; (b) under- and over-triage rates; (c) speed, reproducibility, and interrater reliability; and (d) applicability to special populations (e.g., pediatrics). By mapping strengths and gaps across legacy and next-generation systems, the review aims to inform protocol selection, training priorities, and future tool development under MUCC-aligned, evidence-based frameworks.

Aim and Objectives

This systematic review aims to appraise the diagnostic accuracy and operational effectiveness of prehospital triage tools used in disasters and mass-casualty incidents (MCIs). Specifically, we will (1) estimate sensitivity/specificity for identifying patients needing life-saving interventions (LSIs) and other critical outcomes; (2) summarize under- and over-triage rates; (3) evaluate speed, usability, and inter-rater reliability where available; and (4) compare performance across adult and pediatric tools, legacy algorithms (e.g., START, SALT) and newer tools (e.g., MPTT-24, MITT). (Franc, et al., 2022).

Methods

Design and registration

Systematic review of diagnostic test accuracy (DTA) and effectiveness studies following PRISMA-DTA guidance. Where meta-analysis is feasible, we will use hierarchical/bivariate models for sensitivity/specificity with 95% CIs; otherwise, we will synthesize narratively. Certainty of evidence will be rated with GRADE for test accuracy. (Protocol text—register in PROSPERO prior to extraction.)

Eligibility criteria (PICOS)

- **Population:** Patients (any age) triaged in prehospital or immediate receiving areas during disasters/MCIs, including large exercises or registry re-triage simulating MCI conditions.
- **Index tests:** Any prehospital MCI triage tool (e.g., START, SALT, CareFlight, Sieve/Sort, Sacco, JumpSTART, PTT, SPTT, MPTT/MPTT-24, MITT, Ten-Second Triage).
- **Comparators:** Another triage tool or reference standard.
- **Outcomes (primary):** Correct identification of need for LSI (composite functional gold standard, per MUCC-aligned definitions); under- and over-triage. **Secondary:** mortality, major injury (e.g., ISS≥15), time-to-triage, inter-rater reliability, usability.
- **Study designs:** Prospective/retrospective cohorts, registry re-triage, simulation/live exercises, and guideline-linked implementation studies. Exclude editorials and hospital-only non-MCI triage (e.g., ESI without an MCI context).

Information sources and search strategy

We will search MEDLINE (PubMed), Embase, Web of Science, Scopus, Cochrane, and CINAHL from 2000 to present, plus gray literature (NHS England MITT/TST documents; CDC/EMS.gov MUCC; HHS CHEMM/REMM). Example PubMed string:

("mass casualty" OR disaster* OR "major incident*") AND (triage OR "triage tool*" OR START OR SALT OR MPTT OR MITT OR JumpSTART OR "Ten Second Triage" OR Sieve OR "Triage Sort") AND (sensitiv* OR specific* OR "life-saving intervention*" OR LSI OR accuracy)

Study selection and data extraction

Two reviewers were independently screen titles/abstracts and full texts, extract data (setting, tool, reference standard, thresholds, speed, training, outcomes), and resolve disagreements by consensus/third reviewer.

Risk of bias and applicability

We used **QUADAS-2** for DTA studies and **ROBINS-I** for non-randomized effectiveness studies. Applicability will consider disaster type, responder mix, and pediatric vs adult cohorts.

Data synthesis and analysis

- **Primary analysis:** Sensitivity/specificity for LSI, with HSROC plots where ≥4 comparable studies exist.
- **Secondary:** Pooled under-/over-triage (random-effects proportions) where consistent definitions allow; time-to-triage and inter-rater reliability summarized narratively.
- **Subgroups:** Adult vs pediatric; legacy vs physiology-based tools; real-world vs simulation; region; incident type.
- **Sensitivity analyses:** Excluding high-risk-of-bias studies; varying LSI definitions per consensus lists.

Results (evidence to date — concise synthesis)

Legacy tools: A 2022 meta-analysis found that START lacks sufficient diagnostic accuracy for reliable disaster triage, reinforcing earlier concerns about wide performance variability (Franc, et al., 2022). SALT remains the only MUCC-aligned, nationally harmonizing framework in the U.S., but formal accuracy evidence is mixed and often simulation-based.

Physiology-based next-generation tools: The Modified Physiological Triage Tool and its refinement (MPTT-24) consistently outperform older UK tools (Sieve/Sort) for identifying LSIs (substantially higher sensitivity, with trade-off in specificity). These findings informed NHS England's 2023 adoption of a unified **Major Incident Triage Tool (MITT)** across responders, alongside **Ten-Second Triage (TST)** for very early, rapid sorting.

Pediatrics: Evidence for JumpSTART is limited and inconsistent; a systematic review reported insufficient validation of JumpSTART as a universal pediatric tool. Emerging analyses suggest **SPTT** and **MPTT-24** (age-inclusive with pediatric parameters) achieve higher sensitivity for LSI than JumpSTART or PTT.

Operational targets: In trauma-system benchmarks often referenced in disaster planning, accepted under-triage is ~0–5% and over-triage ~25–35%—levels many studies exceed in field or exercise settings, highlighting ongoing gaps.

Training and technology: National frameworks (MUCC in the U.S.; MITT/TST in England) are pushing standardization. Immersive XR/VR training shows promising gains in triage accuracy, time, and user acceptance, supporting scale-up for preparedness, though more controlled evaluations are needed.

Discussion

Across diverse settings, the central trade-off is unchanged: maximizing sensitivity for time-critical injury while containing over-triage. Evidence synthesized to date indicates that (1) legacy algorithms like START under-perform for LSI discrimination; (2) physiology-driven tools derived against LSI outcomes (MPTT-24) and their implementations (MITT) improve early identification at acceptable speed and complexity; and (3) pediatric triage remains the weakest link, where SPTT/MPTT-24 outperform JumpSTART in retrospective and registry-based analyses but still require prospective external validation during real incidents. PubMed+1england.nhs.ukPMC

For systems design, adopting MUCC-aligned elements (global sort, limited LSIs, simple thresholds) and explicitly using LSI as the reference standard enhances consistency and evidence coherence. Implementation should pair tool selection with frequent, realistic training—now increasingly viable via VR/XR—plus clear operational targets (e.g., minimize under-triage, accept measured over-triage), and periodic audit using real incidents and large-scale exercises.

Practical Implications

- **Tool choice:** Prefer LSI-anchored, high-sensitivity tools (e.g., MPTT-24/MITT) for primary triage; use SALT to maintain MUCC alignment and interoperability where MITT is not adopted.
- **Pediatrics:** Consider SPTT or age-inclusive MITT approaches; do not rely solely on JumpSTART where stakes are high.
- **Training:** Integrate XR/VR scenarios to maintain speed and consistency under stress; monitor transfer to real-world performance.
- **Quality targets:** Track under-/over-triage against trauma benchmarks and investigate deviations after exercises or incidents.

Limitations of the Evidence Base

Much of the literature uses simulation, registry re-triage, or retrospective datasets rather than prospective disaster data; reference standards vary (mortality, ISS, LSI), and pediatric validation remains sparse. These factors limit pooled estimates and generalizability to chaotic real-world MCIs.

Conclusion

Across diverse disaster contexts, the central operational trade-off—maximizing sensitivity for time-critical injuries while containing over-triage—persists. The balance of current evidence supports replacing or supplementing legacy algorithms with physiology-based, LSI-oriented tools such as MPTT-24 and implementations like MITT, which consistently improve identification of patients requiring life-saving interventions without unacceptable complexity.

System-level adoption should be grounded in MUCC-aligned elements (global sort; a limited set of LSIs; simple, reproducible thresholds) and paired with frequent, realistic training. Emerging XR/VR platforms can standardize practice and stress-test decision-making at scale; however, their real-world impact on patient-centered outcomes needs rigorous evaluation. Programs should track under-/over-triage against explicit targets, audit performance after large exercises or incidents, and iteratively refine protocols.

Pediatric triage remains the weakest evidence area. While SPTT and age-inclusive modern tools appear to outperform JumpSTART on LSI sensitivity in retrospective and simulation studies, prospective validation during true MCIs is limited. Services should therefore adopt pediatric approaches that favor sensitivity, while investing in targeted training and research to reduce uncertainty.

Finally, future studies should prioritize prospective, multi-agency evaluations using consistent LSI-based reference standards, capture inter-rater reliability and time metrics alongside clinical outcomes, and report stratified results (adult vs pediatric; mechanism; incident type). Such data will enable robust meta-analysis, narrow practice variability, and, ultimately, improve survivability and stewardship of scarce resources during MCIs.

References

American College of Surgeons Committee on Trauma. (2022). *National guideline for the field triage of injured patients* (2021 update & materials). https://www.facs.org ACS+2ACS+2

Baetzner, A. S., et al. (2025). Mass casualty incident training in immersive virtual reality. *Frontiers/PMC*. https://pmc.ncbi.nlm.nih.gov/ PMC

Biesboer EA, Pokrzywa CJ, Karam BS, Chen B, Szabo A, Teng BQ, Bernard MD, Bernard A, Chowdhury S, Hayudini AE, Radomski MA, Doris S, Yorkgitis BK, Mull J, Weston BW, Hemmila MR, Tignanelli CJ, de Moya MA, Morris RS. Prospective validation of a hospital triage predictive model to decrease undertriage: an EAST multicenter study. Trauma Surg Acute Care Open. 2024 May 2;9(1):e001280. doi: 10.1136/tsaco-2023-001280. PMID: 38737811; PMCID: PMC11086287.

Franc, J. M., et al. (2022). METASTART: A systematic review and meta-analysis of START diagnostic accuracy for disaster triage. *Prehospital and Disaster Medicine*, 37(1), 1–10. https://doi.org/10.1017/S1049023X21001368 PubMedCambridge University Press & Assessment

Gianola, S., Castellini, G., Biffi, A. *et al.* Accuracy of pre-hospital triage tools for major trauma: a systematic review with meta-analysis and net clinical benefit. *World J Emerg Surg* **16**, 31 (2021). https://doi.org/10.1186/s13017-021-00372-1

HHS/REMM & CHEMM. (n.d.). JumpSTART pediatric triage algorithm. https://remm.hhs.gov/ and https://chemm.hhs.gov/ REMMCHEMM

Lerner, E. B., et al. (2014). A consensus-based gold standard for MUCC triage categories. *Prehospital Emergency Care*, 18(2), 217–223. https://pmc.ncbi.nlm.nih.gov/ PMC

Mayo Clinic. (2021). Mass casualty triage guidelines revised. https://www.mayoclinic.org/ Mayo Clinic

NHS England. (2023a). *Major incident triage tools* (MITT & TST explainer). https://www.england.nhs.uk/ england.nhs.uk

NHS England. (2023b). *NHS Major Incident Triage Tool (MITT)* (flowchart). https://www.england.nhs.uk/ (PDF). england.nhs.uk

NHS England EPRR. (2023c). Letter to ambulance services on MITT & TST implementation. https://www.england.nhs.uk/ (PDF). england.nhs.uk

Newgard, C. D., et al. (2022). National guideline for the field triage of injured patients: 2021 recommendations. *Journal of Trauma and Acute Care Surgery*, 93(6), e49–e60. https://pmc.ncbi.nlm.nih.gov/ PMC

Purwadi, H., Breaden, K., McCloud, C., & Pranata, S. (2021). The SALT and START Triage System for Classifying Patient Acuity Level: A Systematic Review. Nurse Media Journal of Nursing, 11(3), 413-427. https://doi.org/10.14710/nmjn.v11i3.37008

Reichard, A. A., et al. (2014–2020). MUCC national implementation materials. EMS.gov. https://www.ems.gov/ EMS.gov

Romig, L., et al. (2022). JumpSTART triage protocol in disaster pediatric patients: Systematic review. *Prehospital and Disaster Medicine*, 37(2), 1–9. https://pubmed.ncbi.nlm.nih.gov/35109962/ PubMed

SALT Working Group. (2008). SALT mass casualty triage: Concept endorsed by ACEP/ACS-COT/NAEMSP et al. *Disaster Medicine and Public Health Preparedness*, 2(4), 245–246. PubMed

StatPearls (Clarkson, L.). (2023). EMS mass casualty triage (overview of START/SALT). NCBI Bookshelf. https://www.ncbi.nlm.nih.gov/books/ المركز الوطني للتكنولوجيا الحيوية

Stéfani GM, de Melo ME, Zardeto HN, Costa VSLP, Lima FS, Cola M. JumpSTART Triage Protocol in Disaster Pediatric Patients: A Systematic Literature Review. Prehospital and Disaster Medicine. 2022;37(2):240-246. doi:10.1017/S1049023X22000127

Vassallo, J., et al. (2017). Derivation/comparative analysis of the MPTT. *Injury*, 48(5), 992–999. https://pubmed.ncbi.nlm.nih.gov/28131484/ PubMed

Vassallo, J., et al. (2018/2020). MPTT-24 development/validation and performance vs Sieve/Sort. *Journal of the Royal Army Medical Corps*, 166(1), 33–37; *Emergency Medicine Journal*, 35(11), 638–644. BMJ Military Health+1White Rose Research Online

Vassallo, J., Moran, C., Cowburn, P., et al. (2022). New NHS prehospital MITT (from MIMMS to MITT). *Emergency Medicine Journal*, 39(11), 800–802. https://emj.bmj.com/ BMJ Evidence-Based Medicine

Vassallo, J., et al. (2021). Comparative analysis of pediatric triage tools (SPTT vs JumpSTART/PTT). *Emergency Medicine Journal*. https://pmc.ncbi.nlm.nih.gov/ PMC

XR/VR effectiveness for MCI triage training: Cardós-Alonso, M. del C., et al. (2024). *International Journal of Emergency Medicine*, 17, 60. https://intjem.biomedcentral.com/ BioMed Central

Zhang J, Liu K, Chen R, Ren J, Wang R, Li L, Jiang J, Hua Y and Wan Y (2023) Assessment of three triage systems by medical undergraduate students using simulated disaster patients: a comparative study. Front. Disaster Emerg. Med. 1:1169851. doi: 10.3389/femer.2023.1169851